Texture Classification and Segmentation Using Wavelet Frames

Michael Unser, *Senior Member, IEEE*

Abstract—**This paper describes a new approach to the characterization of texture properties at multiple scales using the wavelet transform. The analysis uses an overcomplete wavelet decomposition, which yields a description that is translation invariant. It is shown that this representation constitutes a tight frame of l_2 and that it has a fast iterative algorithm. A texture is characterized by a set of channel variances estimated at the output of the corresponding filter bank. Classification experiments with 12 Brodatz textures indicate that the discrete wavelet frame (DWF) approach is superior to a standard (critically sampled) wavelet transform feature extraction. These results also suggest that this approach should perform better than most traditional single resolution techniques (co-occurrences, local linear transform, and the like). A detailed comparison of the classification performance of various orthogonal and biorthogonal wavelet transforms is also provided. Finally, the DWF feature extraction technique is incorporated into a simple multicomponent texture segmentation algorithm, and some illustrative examples are presented.**

I. Introduction

MOST traditional statistical approaches to texture analysis (e.g., co-occurrence matrices [1], second-order statistics [2], Gauss–Markov random fields [3], local linear transforms [4], [5], etc.) are restricted to the analysis of spatial interactions over relatively small neighborhoods [6], [7]. As a consequence, their performance is best for the analysis and segmentation of the class of so-called microtextures. Recent models for human and mammalian vision suggest the existence of an internal spatial/frequency representation that is capable of preserving both local and global information [8]–[11]. These findings have been the basis for several approaches to texture using banks of Gabor filters with different scale and orientation tuning [12]–[15]; these templates are obtained from the modulation of a Gaussian function and are therefore well localized in space and frequency. However, there has been no study providing a clear demonstration of the superiority of these approaches over the more traditional ones.

A potential disadvantage of decompositions of the latter type is that they are computationally quite intensive, especially for the evaluation of low-frequency components. In addition, the outputs of Gabor filter banks are not mutually orthogonal, which may result in a significant correlation between texture features. Finally, these transformations are usually not reversible, which limits their applicability for texture synthesis. Most of these problems can be avoided if one uses the wavelet transform, which provides a precise and unifying framework for the analysis and characterization of a signal at different scales [16]–[19]. The use of a pyramid-structured wavelet transform for texture analysis was first suggested in the pioneering work of Mallat [19]. This initial proposal has been followed by several studies on texture classification with a particular attention to the use of wavelet packets [20], [21], which constitute a multiband extension of the pyramid-structured wavelet transform.

In this paper, a variation of the discrete wavelet transform is introduced for characterizing texture properties. This technique is applied to the problems of texture classification and segmentation. The present analysis method, which is described in Section II, uses an overcomplete wavelet decomposition (the discrete wavelet frame (DWF)) in which the output of the filter banks is not subsampled. Unlike other wavelet-based approaches, this should result in a texture description invariant with respect to translations of the input signal. This property, which appears to be quite desirable in the present context, should yield a better estimation of texture statistics and a more detailed texture characterization at region boundaries. In addition, the oversampled wavelet transform can be viewed as a direct multiscale extension of the local linear transform approach, which is a method that is well understood and that compares favorably with most traditional approaches [5].

Another crucial point is the experimental evaluation of wavelet transform feature extraction techniques and the comparison with other methods. This issue is addressed in Section III, which presents a comparison of classification performance for several orthogonal and semiorthogonal spline wavelet transforms [19], [22]. The influence of the regularity parameter (number of vanishing moments) is also investigated. A representation of special interest is the Gabor-like B-spline (resp. D-spline) wavelet transform for which the synthesis (resp. analysis) filters provide very close approximations of modulated Gaussians [23]. Finally, a simple way of integrating wavelet-based feature extraction into a multicomponent texture segmentation algorithm is described in Section IV.

II. Wavelet Transforms and Frames

The wavelet transform is usually described as a multiresolution decomposition for finite energy functions f of the continuous variable x (i.e., $f(x) \in L_2$) [16], 19], [24]. In the present context, it is more appropriate to consider wavelet representations for discrete signals in l_2 (the space

Manuscript received December 3, 1992; revised March 5, 1995. The associate editor coordinating the review of this paper and approving it for publication was Prof. Robert M. Haralick.

The author is with the Biomedical Engineering and Instrumentation Program, National Center for Research Resources, National Institutes of Health, Bethesda, MD 20892-5766 USA.

IEEE Log Number 9414597.

of square summable sequences), which was a formulation initially proposed by Rioul [25]. This section starts with a definition of the underlying filter banks and then provides a brief review of the main properties of the orthogonal wavelet transform in l_2. In Section II-C, this approach is extended to the case of an overcomplete representation, and the concept of an orthogonal (or tight) wavelet frame is introduced. Finally, this decomposition is related to the method of local linear transforms, which leads naturally to a characterization of textures in terms of simple channel statistics (histograms and variances).

A. Filter Bank Specification and Properties

The purpose of this section is to characterize the perfect reconstruction filter banks underlying the definition of the various types of wavelet transforms considered here. Such systems are entirely specified in terms of a prototype filter h that satisfies the standard quadrature mirror filter condition

$$H(z)H(z^{-1}) + H(-z)H(-z^{-1}) = 1 \tag{1}$$

where $H(z)$ denotes the z-transform of h. The filter h is also required to satisfy the lowpass constraint $H(z)|_{z=1} = 1$. A complementary highpass filter is obtained by shift and modulation

$$G(z) = zH(-z^{-1}). \tag{2}$$

These prototypes are then used to generate, in an iterative fashion, a sequence of filters of increasing width (indexed by i)

$$H_{i+1}(z) = H(z^{2^i})H_i(z) \tag{3}$$

$$G_{i+1}(z) = G(z^{2^i})H_i(z), \quad (i = 0, \ldots, I-1) \tag{4}$$

with the initial condition $H_0(z) = 1$. Equivalently, in the signal domain, we have the two-scale relation

$$\begin{cases} h_{i+1}(k) = [h]_{\uparrow 2^i} * h_i(k) \\ g_{i+1}(k) = [g]_{\uparrow 2^i} * h_i(k) \end{cases} \tag{5}$$

where the notation $[\cdot]_{\uparrow m}$ denotes the upsampling by a factor of m. The effect of one iteration is more or less to dilate the filters h_i and g_i by a factor of two.[1] Such a sequence of filters can therefore be used to decompose a signal in subbands of approximately one octave each. The properties of such a decomposition are best understood by analyzing the multiband characteristics of the underlying filter bank; some examples are shown in Fig. 1. It is not difficult to verify that such sequences of filters also satisfy the identity

$$|H_i(e^{j2\pi f})|^2 + \sum_{k=1}^{i} |G_k(e^{j2\pi f})|^2 = 1 \tag{6}$$

and therefore provide a full coverage of the frequency domain.

[1] In the continuous formulation, it is precisely a dilation by factor of two. In the discrete case, it is only approximatively the case unless the filter h satisfies an additional interpolation condition.

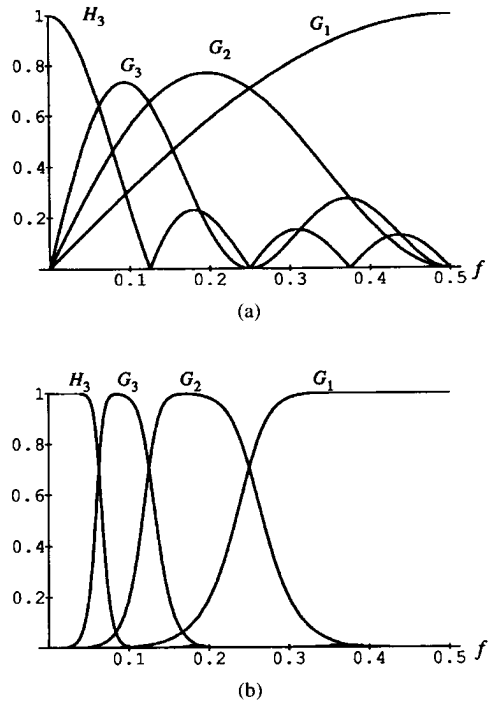

Fig. 1. Frequency responses of the discrete filter bank associated with the Battle/Lemarié spline wavelet transform for $n = 0$ (Haar) and $n = 3$ (cubic splines).

B. Orthogonal Wavelet Bases of l_2

The sequence of discrete filters defined above can be used to construct an orthogonal wavelet decomposition of l_2 [25]. For this purpose, one defines the discrete normalized basis functions

$$\varphi_{i,l}(k) = 2^{i/2} h_i(k - 2^i l) \tag{7}$$

$$\psi_{i,l}(k) = 2^{i/2} g_i(k - 2^i l) \tag{8}$$

where i and l are the scale and translation indices, respectively; the factor $2^{i/2}$ is an inner product normalization. We then consider the sequence of nested subspaces $l_2 = V_0 \supset V_1 \supset V_2 \supset \cdots \supset V_I$, where $V_i = \text{span}\{\varphi_{i,l}\}_{l \in Z}$ is the approximation space at resolution i. Note that the embedding of these subspaces comes as a direct consequence of the two scale relations (5). We also introduce the subspaces W_i, $(i = 1, \ldots, I)$, where W_i is the residue (or detail) space at resolution i and is defined as the orthogonal complement of V_i with respect to V_{i-1} (i.e., $V_{i-1} = V_i + W_i$ and $V_i \perp W_i$). It can be shown by induction that the families of sequences $\{\varphi_{i,l}\}_{l \in Z}$ and $\{\psi_{i,l}\}_{l \in Z}$ provide orthonormal bases of V_i and W_i, respectively. We can therefore evaluate the coordinates of the orthogonal projection of a discrete signal $x \in l_2$ in the spaces V_i and W_i by forming the inner product with the corresponding basis functions. Specifically, the minimum l_2-norm approximation of x at scale i (orthogonal projection into V_i) is given by

$$x_{(i)}(k) = \sum_{l \in Z} s_{(i)}(l) \varphi_{i,l} \tag{9}$$

$$s_{(i)}(l) = \langle x(k), \varphi_{i,l}(k) \rangle_{l_2} \tag{10}$$

where $\langle \cdot, \cdot \rangle_{l_2}$ denotes the standard l_2 inner product and where $\varphi_{i,0}(k) = 2^{i/2}h_i(k)$ is the discrete scaling function at resolution i. Likewise, the corresponding residue, which is the projection of x into W_i, is given by the complementary wavelet expansion

$$x_{(i-i)}(k) - x_{(i)}(k) = \sum_{l \in Z} d_{(i)}(l)\psi_{i,l} \tag{11}$$

$$d_{(i)}(l) = \langle x(k), \psi_{i,l}(k) \rangle_{l_2} \tag{12}$$

where $\psi_{i,0}(k) = 2^{i/2}g_i(k)$ is the discrete wavelet at scale i. Finally, by combining the residues over all scales down to a given depth I, we obtain the full discrete wavelet expansion of a signal $x \in l_2$

$$x(k) = \sum_{l \in z} s_{(I)}(l)\varphi_{I,l} + \sum_{i=1}^{I} \sum_{l \in Z} d_{(i)}(l)\psi_{i,l} \tag{13}$$

where the $d_{(i)}$'s are the wavelet coefficients and where the $s_{(I)}$'s are the expansion coefficients of the coarser signal approximation $x_{(I)}$ (cf. (9)). It also follows from this construction that the family of sequences $\{\varphi_{I,l}, \psi_{1,l}, \psi_{2,l}, \ldots, \psi_{I,l}\}_{l \in Z}$ constitutes an orthonormal basis of l_2.

What is important in the context of texture analysis is that the analysis formulas (10) and (12) can be interpreted in terms of simple filtering and down-sampling operations. Specifically, we can write that

$$\begin{cases} s_{(I)}(l) = 2^{I/2}[h_I^T * x]_{\downarrow 2^I}(l) \\ (d)_i(l) = 2^{i/2}[g_i^T * x]_{\downarrow 2^i}(l), \quad (i = 1, \ldots, I) \end{cases} \tag{14}$$

where the symbol ^{T} denotes the reversal (or transpose) operator (i.e, $h^T(k) = h(-k)$) and where $[\cdot]_{\downarrow m}$ is the downsampling by a factor of m. However, this approach is not the one that is usually used for computing the discrete wavelet transform. There is a much more efficient algorithm that uses the filters h and g directly and performs the decomposition using a critically sampled tree-structured quadrature mirror filter bank [19].

Finally, note that there is a direct equivalence between wavelet transforms and perfect reconstruction filter banks. It is possible, for instance, to drop the quadrature mirror filter requirement (1) and extend the construction to the biorthogonal case [17], [26]. This approach offers greater flexibility in the choice of the basis functions at the expense of orthogonality. An example worth mentioning is the B-spline wavelet transform, which has the property that the basis functions are very similar to Gabor functions [23]. This representation provides a link between the wavelet formulation and some of the Gabor transform approaches to texture mentioned in the introduction.

C. Wavelet Frames of l_2

As already pointed out in the introduction, a simple integer shift of the input signal will usually result in a nontrivial modification of the discrete wavelet transform, as defined in Section II-B. As far as feature extraction is concerned, this behavior is inadequate, for one usually thinks of "texture" as a translation-invariant (or stationary) property. A natural way to overcome this limitation is to perform an analysis of the input signal in terms of the overcomplete family of templates

$$S = \{g_1(k-l), \ldots, g_I(k-l), h_I(k-l)\}_{l \in Z}$$

which is similar to computing the discrete wavelet transform for all possible integer shifts of the input signal. This approach leads to the simple decomposition algorithm

$$\begin{cases} d_i^{\mathrm{DWF}}(k) := \langle g_i(k-l), x(k) \rangle_{l_2} = g_i^T * x(k), \\ \qquad\qquad\qquad\qquad\qquad (i = 1, \ldots, I) \\ S_I^{\mathrm{DWF}}(k) := \langle h_I(k-1), x(k) \rangle_{l_2} = h_I^T * x(k) \end{cases} \tag{15}$$

which is the nonsampled version of (14). From now on, we will primarily consider such redundant representations and drop the superscript "DWF," which has been used here to distinguish these values from the wavelet coefficients in the previous section. Because of the special structure of the analysis filter bank, this decomposition has a number a remarkable properties that are associated with the mathematical concept of a frame [27], [28]. These include energy conservation and a particularly simple reconstruction algorithm.

Proposition 1: The family of sequences S is a tight frame of the Hilbert space l_2.

Proof: To prove that S constitutes a frame of l_2, we need to show that there exist two constants A and B such that

$$A \cdot \|x\|_{l_2}^2 \le \sum_{l \in Z} \langle x(k), h_I(k-l) \rangle^2 + \sum_{i=1}^{I} \sum_{l \in Z} \langle x(k), g_i(k-l) \rangle^2 \le B \cdot \|x\|_{l_2}^2. \tag{16}$$

We start by using Parseval's formula to compute the energies in the different channels

$$\|d_i\|_{l_2}^2 = \int_0^1 |G_i(e^{j2\pi f})|^2 |X(f)|^2 df$$

$$\|s_I\|_{l_2}^2 = \int_0^1 |H_I(e^{j2\pi f})|^2 |X(f)|^2 df$$

where $X(f)$ denotes the Fourier transform of the input signal. We then sum the individual terms and use (6) to show that we have the energy conservation property

$$\|x\|_{l_2}^2 = \int_0^1 |X(f)|^2 df = |s_I\|_{l_2}^2 + \sum_{i=1}^{I} \|d_i\|_{l_2}^2. \tag{17}$$

By definition, $s_i(l) = \langle x(k), h_i(k-l) \rangle$ and $d_i(l) = \langle x(k), g_i(k-l) \rangle$. Therefore, we can rewrite this last formula in the form of inequality (16) with the simple bounds $A = B = 1$, which also proves that the frame is tight. □

A remarkable property of tight frames in general is that they lead to a representation of a signal that is analogous to that associated with an orthogonal basis. Specifically, if $S = \{\varphi_i\}_{i \in Z}$ is a tight frame of a Hilbert space H, then

$$\forall x \in H, \quad x = \frac{1}{A} \sum_{i \in Z} \langle x, \varphi_i \rangle_H \, \varphi_i \tag{18}$$

where $\langle \cdot, \cdot \rangle_H$ is the corresponding inner product [28]; the fundamental difference with an orthogonal system is that the

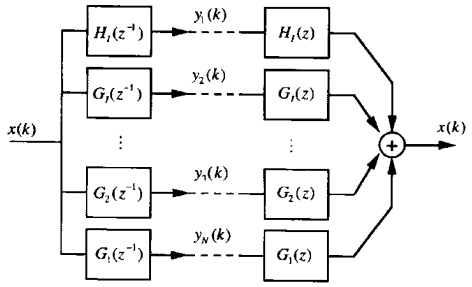

Fig. 2. Signal analysis and synthesis using the discrete wavelet frame decomposition.

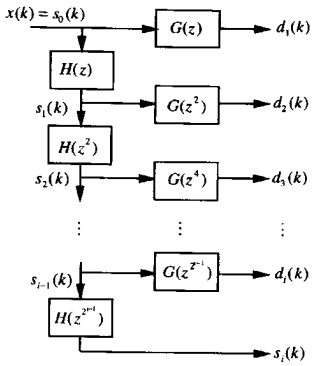

Fig. 3. Fast iterative implementation of the discrete wavelet frame decomposition.

representation may be redundant. In the present context where $H = l_2$, this property, together with the definition of the wavelet coefficients in (15), leads to the simple reconstruction formula

$$x(k) = \sum_{l\in Z} s_I(l)h_I(k-l) + \sum_{i=1}^{I}\sum_{l\in Z} d_i(l)g_i(k-l) \quad (19)$$

which can also be interpreted in terms of a reconstruction filter bank

$$x(k) = h_I * s_I(k) + \sum_{i=1}^{I} g_i * d_i(k). \quad (20)$$

The principle of this decomposition is illustrated in Fig. 2. The fact that the whole system acts as the identity operator also follows directly from (6).

To simplify the implementation, we take advantage of the two-scale relation (5) and obtain a fast iterative decomposition algorithm

$$\begin{cases} s_{i+1}(k) = [h]_{\uparrow 2^i} * s_i(k) \\ d_{i+1}(k) = [g]_{\uparrow 2^i} * s_i(k) \end{cases}, \quad (i = 0, \ldots, I) \quad (21)$$

with the initial condition $s_0 = x$ (cf. block diagram in Fig. 3). Each step involves a convolution with the basic filters h and g, which are expanded by insertion of an appropriate number of zeros between taps. A similar procedure can also be used for the implementation of the reconstruction formula (20). Note that the complexity of this type of algorithm is the same for all iterations. It is simply proportional to the number of samples.

Such an overcomplete analysis can also be extended to the biorthogonal case. In this more general situation, one looses the energy conservation (or tight frame) property but one has more freedom in the choice of the filters H and G. It is not difficult to show that the frame condition (16) can be translated into a simpler constraint on the frequency response of these filters

$$0 < m \le |H(e^{j2\pi f})|^2 + |G(e^{j2\pi f})|^2 \le M < +\infty \quad (22)$$

where m and M are two positive constants. This relation guarantees the reversibility of the decomposition. In fact, we obtain a reconstruction algorithm that is almost identical to (20) if we can find two biorthogonal filters $\mathring{h}$ and $\mathring{g}$ such that

$$H(z)\mathring{H}(z^{-1}) + G(z)\mathring{G}(z^{-1}) = 1. \quad (23)$$

D. Texture Characterization

To extend these approaches in higher dimensions, we use a standard tensor product formulation [19]. In two dimensions, there will be four distinct types of basis functions (or filters) corresponding to the different crossproducts of the 1-D functions φ and ψ. The decomposition can thus be obtained by successive 1-D processing along the rows and columns of an image. The same system can also be used for the analysis of signals that are realizations of wide-sense stationary processes. For this purpose, however, it is necessary to replace the previous notion of the squared l_2-norm by the average power $E\{x^2\}$ (or variance if the signal is zero mean) and modify all energy formulas accordingly.

The filter bank analysis system specified by (15) can be viewed as a special case of the local linear transform method [5]. In this approach, we rearrange the output of the filter bank into the N-component vector, where N is the number of subbands

$$\begin{aligned} \mathbf{y}(k,l) &= (y_i(k,l))_{i=1,\ldots,N} \\ &= [s_I(k,l) \quad d_I(k,l) \quad \cdots \quad d_1(k,l)]^T \end{aligned} \quad (24)$$

and interpret the result of the analysis for a given spatial index (k,l) as a local linear transformation of the input vector $\mathbf{x}(k,l)$; the latter is simply a block representation of the input image centered on the current position. A 2-D separable wavelet transform with a depth I typically yields $N = 1 + 3I$ such feature channels. The texture is then characterized by the set of N first-order probability density functions $p(y_i)$, $i = 1, \ldots N$. Alternatively, we can get a more compact representation in terms of the channel variances $\mathrm{Var}\{y_i\}$. A statistical justification for this approach can be found in [5]. The success of this method obviously depends on the judicious selection of the filter bank.

In practice, the channel variances are estimated from the average sum of squares over a region of interest R of the given texture type

$$v_i = \frac{1}{\#R}\sum_{(k,l)\in R} y_i^2(k,l) \quad (25)$$

where $\#R$ denotes the number of pixels in R. Note that the lowpass condition $H(z)|_{z=1} = 1$ implies that $E\{y_1\} = E\{x\}$, and $E\{y_i\} = 0$, for $i = 2, \ldots, N$. It may therefore

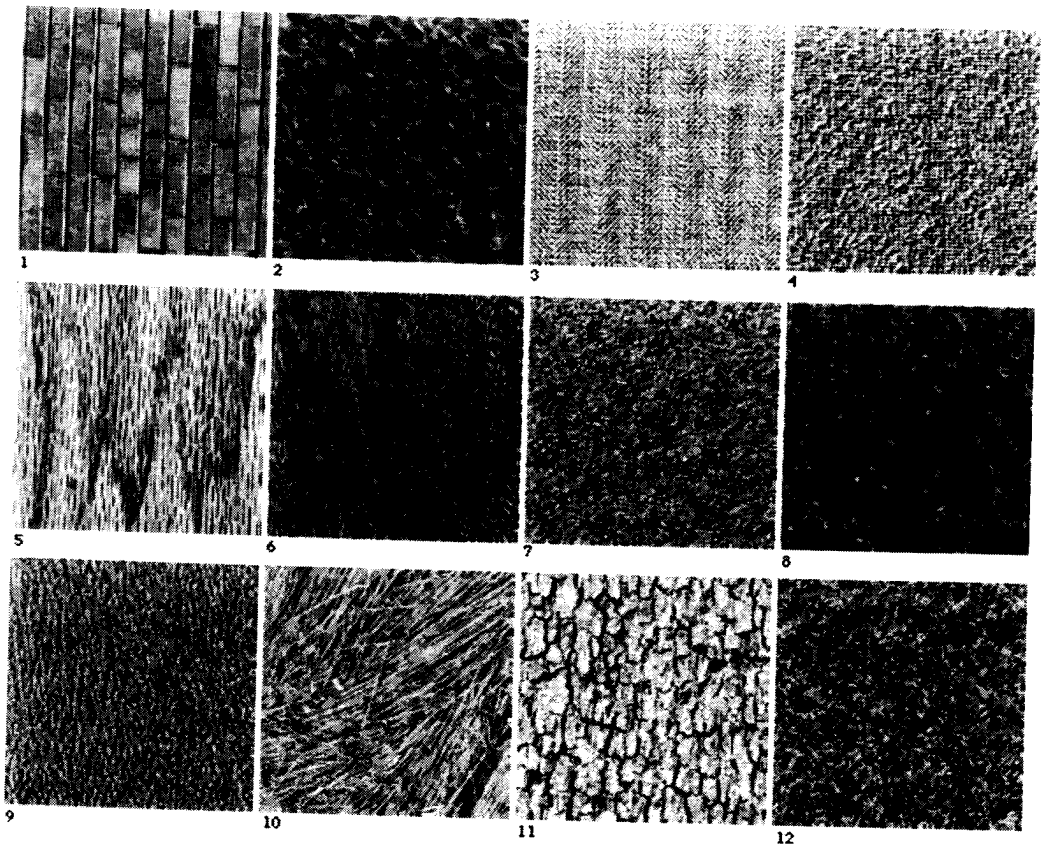

Fig. 4. Twelve Brodatz textures used for classification experiments.

be advisable to adjust the feature computed in the lowpass channel by subtracting $E\{x\}^2$ to get a better variance estimate.

If the analysis is performed through the discrete wavelet transform, there will be fewer coefficients in a given channel because of subsampling. We can still obtain variance estimates from the average sum of squares of the available wavelet coefficients in the region of interest. The expected values of the texture features will be the same as in the previous case (up to a proportionality factor), but their variability will be increased with a possible adverse effect on classification performance.

III. Texture Classification

A. Experimental Data Set

We performed classification experiments using the 12 Brodatz textures displayed in Fig. 4. The original photographs were digitized and converted into 256 × 256 image arrays with 256 levels of gray. The spatial uniformity of these images was improved by using a local standardization procedure with local mean and variance estimates computed over a 65 × 65 sliding window. In order to make the classification task more difficult, the images had their histogram equalized with a requantization to 32 levels. The experimental data were therefore indistinguishable on the basis of first-order statistics only. The wavelet and filter bank decompositions were performed by processing the individual images globally. For each texture, a total of 64 independent feature vectors $\mathbf{v} = (v_1, \ldots, v_N)$ was evaluated, using (25) over a series of 32 × 32 nonoverlapping subregions.

B. Classification Algorithm

The class conditional probability density functions were assumed to be multivariate Gaussian with mean vectors and covariance matrices $(\mathbf{m}_k, \mathbf{C}_k), (k = 1, \ldots, K)$, where $K = 12$ is the total number of classes. Under such an assumption, the minimum error Bayes classifier is equivalent to assigning a texture sample with feature vector $\mathbf{v}$ to the class with minimum distance value

$$d_k(\mathbf{v}) = (\mathbf{v} - \mathbf{m}_k)^T \mathbf{C}_k^{-1} (\mathbf{v} - \mathbf{m}_k) + \log(\det(\mathbf{C}_k)). \quad (26)$$

For each pattern tested, the training was performed on the remaining set (leaving-one-out method) using the maximum likelihood estimates of the distribution parameters.

C. Classification Results

In a first series of experiments, we chose to compare the performance of the discrete wavelet transform (DWT) and wavelet frame (DWF) approaches. For this purpose, we considered the series of quadrature mirror filters associated with the Battle–Lemarié orthogonal spline wavelet transforms. This family is indexed by the order of the spline n, which also characterizes the regularity of the underlying continuous basis functions. This choice was made for the following reasons. First, the underlying (discrete or continuous) basis functions are symmetrical, which means that there is no phase distortion

TABLE I
PERCENT OF CORRECT CLASSIFICATION FOR VARIOUS ORTHOGONAL DISCRETE WAVELET TRANSFORMS (DWT'S) AND WAVELET FRAMES (DWF'S) AS A FUNCTION OF THE ORDER n AND THE NUMBER OF SCALES I

Type of decomposition	n=0	n=1	n=3
1 scale (4 features)			
DWT	91.93 %	91.80 %	86.72 %
DWF	96.48 %	93.88 %	91.54 %
2 scales (7 features)			
DWT	95.44 %	98.18 %	97.92 %
DWF	98.83 %	98.70 %	98.44 %
3 scales (10 features)			
DWT	96.35 %	98.44 %	98.31 %
DWF	99.35 %	99.22 %	99.22 %

and that the spatial localization of the wavelet coefficients is well preserved. We feel that this property is extremely relevant for our purpose, especially for texture segmentation. Second, this family includes two important limiting cases. For $n = 0$, we have the Haar transform with the simplest possible filter $H^0(z)(1 + z^{-1})/2$. At the other extreme, for $n \rightarrow +\infty$, we obtain the bandlimited sinc-wavelet, which corresponds to the ideal lowpass filter $h^\infty(k) = \text{sinc}(k/2)$. Increasing n will result in a progressive spatial delocalization of the basis function, but the benefit will be a decrease in the correlation between the various branches of the filter bank.

The classification results for $n = 0$, 1, and 3 are given in Table I. The first observation is that the DWF method always outperforms the DWT, which is consistent with our expectation. In some cases, the improvement is quite substantial. It is also clear that a true multiresolution feature extraction with two or three levels is preferable to a local analysis with one level only. In fact, the DWF with $n = 0$ and $I = 1$ is equivalent to the local linear transform (LLT) method using the 2×2 Hadamard transform described in [5]. For comparison, the LLT using the nine-channel variances associated with the 3×3 discrete sine transform yields 99.22% correct classification. Those results are mentioned because the LLT usually compares quite favorably with most other standard statistical methods (co-occurrence, correlation, etc.) and can therefore be used as the reference method for a single resolution analysis A detailed comparison can be found in [5]. Finally, note that increasing the regularity of the filter bank does not necessarily imply better classification performance. In fact, the results obtained with the DWF for the simplest case $n = 0$ are surprisingly good.

A second series of experiments was conducted to compare the performance of various orthogonal and nonorthogonal wavelet transforms. We were especially interested in determining the properties of the representation (space-frequency localization, orthogonality, regularity, etc.) that are regarded to be the most relevant for texture discrimination. We chose to compare the performance of several members of the class of polynomial spline wavelet transforms [22]. For a fixed value of n, these DWT's are equivalent in the sense that they use the same underlying approximation subspaces, but they differ in the shape and properties of the basis functions.

TABLE II
PERCENT OF CORRECT CLASSIFICATION FOR VARIOUS DISCRETE WAVELET TRANSFORMS AS A FUNCTION OF THE ORDER n AND THE NUMBER OF SCALES I

Wavelet transform	n=0	n=1	n=3
1 scale (4 features)			
Battle/Lemarié	91.93 %	91.80 %	86.72 %
B-spline	-	87.89 %	76.30 %
D-spline	-	91.54 %	87.63 %
2 scales (7 features)			
Battle/Lemarié	95.44 %	98.18 %	97.92 %
B-spline	-	95.18 %	85.55 %
D-spline	-	97.79 %	96.35 %
3 scales (10 features)			
Battle/Lemarié	96.35 %	98.44 %	98.31 %
B-spline	-	95.57 %	87.76 %
D-spline	-	99.09 %	96.88 %

In addition to the orthogonal Battle–Lemarié transform, we considered the B-spline and D-spline wavelet transforms. The B-spline transform was chosen because the underlying basis functions are very close to Gabor functions and therefore optimally localized in space and frequency [23]. The biorthogonal analysis filters, on the other hand, have a comparatively poor localization with a much slower decay in space. The D-spline representation is the dual of the B-spline wavelet transform, which means that it is the analysis filter bank that inherits the good space-frequency localization properties. Texture features were extracted using the standard fast wavelet algorithm that produces a critically sampled representation [22]. The results of classification are summarized in Table II. Here again, the percentage of correct classification improves with the number of scales. On the other hand, increasing the order of the splines does not necessarily have a beneficial effect. The behavior of the orthogonal and D-spline representations is very similar. Both of these transforms clearly perform much better than the B-spline DWT. This suggests that a good space localization of the analysis filter bank is more important for discrimination than the localization of the basis function themselves. If one also takes into account the number of computations, it appears that the most promising approach is the D-spline wavelet transform for which the analysis filter bank is FIR. In practice, we would recommend using either the Haar transform (with $n = 0$) or the D-spline transform with $n = 1$. One should also expect a further improvement (as in Table I) for an analysis based on an overcomplete representation.

Another observation that holds for this whole series of experiments is that classification errors usually occurred between textures pairs that are the most difficult to discriminate visually (see, for example, the confusion matrix in Table III).

IV. TEXTURE SEGMENTATION

The next question is how to modify the previous wavelet-based feature extraction method to make it suitable for texture segmentation. The approach taken here is a variation of the multiresolution feature extraction and selection method developed previously for the LLT [29]. The key idea is

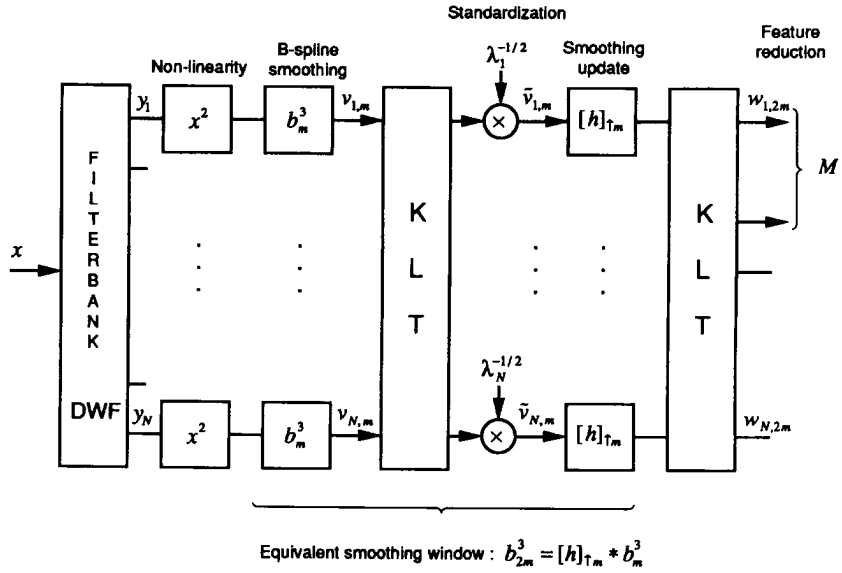

Fig. 5. Block diagram of an image processing system for the wavelet-based segmentation of textured images.

TABLE III
DETAILED CLASSIFICATION RESULTS OBTAINED USING THE HAAR DWF WITH TWO LEVELS

Assigned class →	1	2	3	4	5	6	7	8	9	10	11	12
1	64	0	0	0	0	0	0	0	0	0	0	0
2	0	64	0	0	0	0	0	0	0	0	0	0
3	0	0	64	0	0	0	0	0	0	0	0	0
4	0	0	0	64	0	0	0	0	0	0	0	0
5	0	0	0	0	64	0	0	0	0	0	0	0
6	0	0	0	0	0	64	0	0	0	0	0	0
7	0	0	0	0	0	0	62	0	0	0	0	2
8	0	0	0	0	0	0	0	64	0	0	0	0
9	0	0	0	0	0	0	0	0	64	0	0	0
10	0	0	0	0	0	0	0	0	0	64	0	0
11	0	0	0	0	0	0	0	1	0	0	63	0
12	0	0	0	0	0	0	5	0	0	1	0	58

Number of errors = 9 (out of 768)
Total score = 98.83 %

to use local estimates of the wavelet variances as texture features and to apply an appropriate feature reduction to simplify the segmentation process. This is achieved indirectly by performing an image analysis at two different scales (m) and $(2m)$. The corresponding system is schematically represented in the block diagram in Fig. 5. Note that the segmentation is performed at the coarser scale $(2m)$ after feature reduction; the information at the finer scale m is used merely to improve the estimation of the feature reducing transformation. We will now briefly consider the various components of the system and justify the approach.

A. Feature Extraction

The texture features that are used for segmentation are local energy measures computed over a window centered on the current spatial location. The amount of averaging (size of the estimation window) is specified by the scale parameter m. We have chosen here to use a cubic B-spline window filter that provides a very close approximation of a Gaussian and yet offers the advantage of a fast algorithm whose complexity is independent of the parameter m [30]. Specifically, the ith texture feature at resolution m is computed by convolving the corresponding squared (or rectified) output of the wavelet filter bank

$$v_{i,m}(k,l) = b^3_m * |y_i(k,l)|^2 \tag{27}$$

where b^3_m denotes the separable cubic B-spline kernel enlarged by an integer factor of m. Note that we also need to subtract the mean value in the lowpass channel $(i = 1)$ to obtain a true variance estimate; this correction is not necessary in the wavelet channels $(i > 1)$ because their means are zero by construction. A complete description of the cubic B-spline filter is given in the Appendix. The use of this particular operator can be justified based on the following properties:

i) It is extremely well localized in the sense specified by the uncertainty principle.
ii) It is very close to being isotropic.
iii) It can be implemented efficiently through a cascade of simple 1-D moving average filters that are applied successively along the rows and the columns of the image; the same algorithm is applicable for any value of m.

In our method, we specify the scale parameter m and compute the intermediate (fine resolution) N-component image of local texture features $\mathbf{v}_m(k,l)$ according to (27). Since the cubic B-spline kernel also satisfies the two-scale relation (A5), we then use this property to double the scale of our previous variance estimates through the following update procedure:

$$\mathbf{v}_{2m}(k,l) = [h]_{\uparrow m} * \mathbf{v}_m(k,l) \tag{28}$$

where the convolution is applied to each vector component separately and where the filter $[h]_{\uparrow m}$ is the 2-D separable version of the binomial filter defined by (A6) expanded by a factor of m. The end result of this operation is a set of local variance estimates over a window of size $2m$ ($\sigma_{\text{eq}} = 2m/\sqrt{3}$) specified by the separable 2-D smoothing kernel $b^3_{2m}(k,l) = b^3_{2m}(k) \times b^3_{2m}(l)$; these are the texture features that are used implicitly for segmentation. The additional

information available at the finer scale m is used indirectly to improve the estimation of the feature reducing transformation.

B. Linear Feature Reduction

Assuming that the multicomponent image $\mathbf{v}_m(k,l)$ has a global zero mean, we first evaluate the scatter matrix at resolution m as

$$\boldsymbol{S}_1 = \boldsymbol{S}(\mathbf{v}_m) := \frac{1}{\#R_0} \sum_{(k,l)\in R_0} \mathbf{v}_m(k,l) \cdot \mathbf{v}_m(k,l)^T \quad (29)$$

where R_0 denotes the set of spatial indices within the boundaries of the image to be segmented. The eigenvectors of this matrix are used to specify the Karhunen–Loève transform (KLT), which is then applied to the data and is followed by a standardization. The net effect of this sequence of transformations is to produce an auxiliary feature vector $\tilde{\mathbf{v}}_m(k,l)$ with a whitened scatter matrix, i.e., $\mathbf{S}(\tilde{v}_m) = \mathbf{I}_N$. The next filtering step in Fig. 5 doubles the scale of the averaging window according to (28). Finally, a second KLT is used to rediagonalize the scatter matrix of the transformed feature vector $\mathbf{w}_{2m}(k,l)$ at resolution $2m$.

Since the spatial smoothing is the same in each channel, the resulting vector $\mathbf{w}_{2m}(k,l)$ can also be rewritten as

$$\mathbf{w}_{2m}(k,l) = \mathbf{U} \cdot \mathbf{v}_{2m}(k,l) \quad (30)$$

where $\mathbf{U}$ is an $N \times N$ transformation matrix and where $\mathbf{v}_{2m}$ is given by (28). It is not difficult to show that the whole procedure described in Fig. 5 provides an indirect solution to the generalized eigenvector problem

$$\lambda \mathbf{S}_1 \mathbf{t} = \mathbf{S}_2 \mathbf{t} \quad (31)$$

where $\mathbf{S}_1 = \mathbf{S}(\mathbf{v}_m)$ and $\mathbf{S}_2 = \mathbf{S}(\mathbf{v}_{2m})$. The transformation matrix in (30) can therefore also be expressed as $\mathbf{U} = [\mathbf{t}_1 \ \cdots \ \mathbf{t}_N]^T$, where the $\mathbf{t}_i$'s are the eigenvectors of (31) ranked by order of decreasing eigenvalues. This ordering is also representative of their discriminatory power. Since smoothing decreases variability, the eigenvalues will be smaller than one. The number of significant features (M) can be determined by retaining only those components with an eigenvalue greater than some threshold α (which depends on the amount of smoothing). In any case, this number should be no greater than $K-1$, where K is the number of distinct texture regions. We have found this feature reduction method to be more efficient for segmentation than the standard KLT. Its superiority in this particular application can be attributed to the following properties [29]:

1) Unlike the KLT, it is invariant to any linear transformation of the input vector (such as a change in gain for some of the analysis filters).
2) The transformation will extract the features for which additional spatial smoothing produces the least relative energy reduction. These features should also be those that contain the largest inter-region differences in their mean values.
3) The transformation is equivalent to Fisher's multiple linear discriminants [31] (which is optimal for classification), provided that the two following assumptions are met:
   i) The feature means over the different texture regions are preserved from one resolution to the next;
   ii) the multiple linear discriminant transformation is the same at both resolutions. These conditions are usually well satisfied in practice, at least as long as the smoothing window is smaller than the texture regions.

For this reason, it is not advisable to use too large a value of m (oversmoothing) to estimate the feature reducing transformation. Note that it is always possible *a posteriori* to coarsen the representation by applying additional levels of smoothing to the reduced feature vector $\mathbf{w}_{2m}$ (cf. (28)). Alternatively, one could also apply a multivariate region growing procedure that would merge neighboring feature vectors according to their similarity while attempting to preserve region boundaries.

C. Segmentation by Clustering in Feature Space

The final step of segmentation involves clustering, which can be achieved through the standard K-means algorithm [32]. The main advantage of the present approach is that these computations can be performed in a lower dimensional space that essentially preserves the discriminative information and provides features that are approximately decorrelated (cf. Property (3)). In the simplest case of a two-region segmentation problem, there will be only one significant feature component, and clustering is achieved by simple minimum-error thresholding, as described in [29].

D. Experimental Results

The whole segmentation procedure is illustrated in Fig. 6 using a test image generated by combining two Brodatz textures. As in the previous experiment, the two textures had their histograms equalized and are therefore indistinguishable on the basis of first-order statistics only. The coefficient images of the Haar wavelet frame with a depth of two are shown in Fig. 6(c1–c7). The first four images (c1–c4) represent the coarser components at scale $i = 2$. The first one contains the lowpass information, whereas the channels c2, c3, and c4 represent the midband horizontal, vertical, and diagonal components, respectively. The finer scale information is contained in the channels c5–c7, which corresponds to simple horizontal, vertical, and diagonal edge detectors, respectively. These filtered images were then rectified (or squared) and smoothed individually using the procedure described by (27). The resulting local texture features at resolution $2m = 8$ are shown in Fig. 6(d1–d7). A measure of the discrimination power for each individual feature is provided by the F-statistic (normalized ratio of the inter-region and intra-region variances) computed using the reference map in Fig. 6(b). In this particular example, the features that provide the strongest discrimination are the lowpass (d1) and the horizontal/vertical components (d2) and (d6). The features are then all combined into a single one (Fig. 6(e)) by using the unsupervised two-scale linear feature reduction technique described in Section IV-B with $m = 4$. Clearly, the discrimination power of this new quantity (as measured by the F-statistic) is better than that of any of the initial features on their own. The two texture regions can now be separated by simple thresholding.

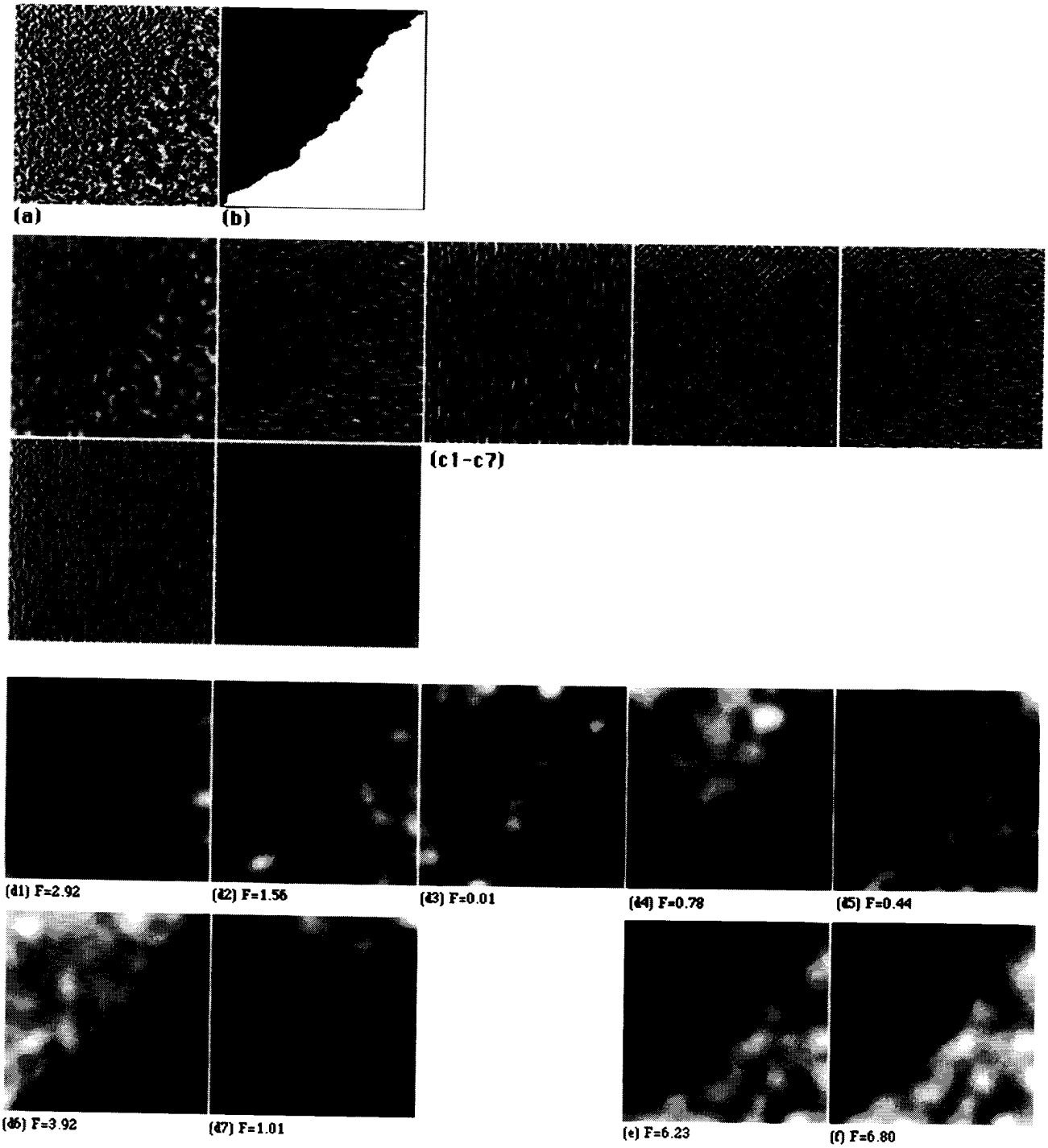

Fig. 6. Illustration of the feature extraction/reduction algorithm: (a) Test image; (b) region labels (gold standard); (c1–c7) filtered channels obtained by applying the Haar wavelet frame decomposition to (a); (d1–d7) smoothed feature planes at scale $m = 8$; (e) most significant rotated feature plane using the two-scale covariance diagonalization algorithm; (f) optimal feature reduction of (d) using Fisher's linear discriminants.

The performance of this feature reduction technique may be compared with the optimal supervised approach (Fisher's linear discriminants), which is given for reference in Fig. 6(f). The present approach is also superior to a standard KLT feature extraction, which in this particular case yields a F-ratio of 5.14. We have also identified other examples of texture pairs for which the improvement of our approach over the KLT is much more significant; this usually happens when the most discriminating wavelet channels have relatively low energy. In addition, there is no guarantee that the components of the KLT will be ordered according to their discriminating power, in constrast to the proposed feature reduction method.

This approach is also applicable for images composed of more than two texture regions. In such a case, the number of significant features (M) is usually greater than one and will require the use of a multidimensional decision procedure

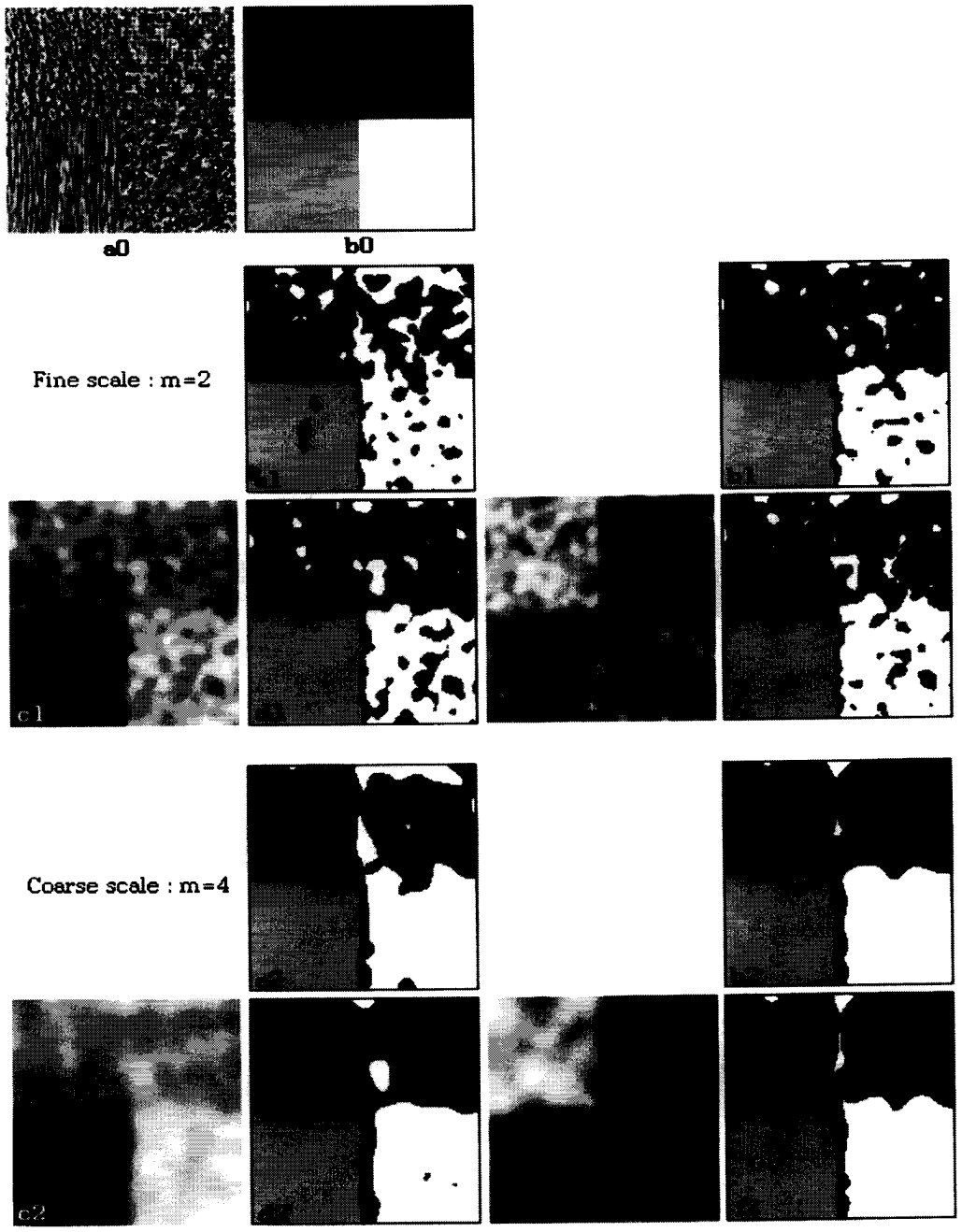

Fig. 7. Examples of textures segementation (fine resolution versus coarse resolution): (a0) 128×128 test image with four texture regions; (b0) region labels (gold standard); (a1–a2) results of full K-means segmentation (Euclidean distance without feature reduction); (b1–b2) full K-means segmentation (weighted distance without feature reduction); (c1–c2) first reduced component; (d1–d2) threshold-based segmentation; (e1–e2) second reduced component; (f1–f2) reduced K-means segmentation (weighted distance using features c and e).

(clustering) more sophisticated than simple thresholding. Fig. 7 presents an example that involves four distinct texture regions with identical (equalized) histograms. In this case, we used a first order Battle–Lemarié DWF feature extraction with a depth of two ($N = 7$). The first set of images shows the results of a fine scale segmentation ($2m = 4$). The results of a full K-means segmentation that is applied to the seven channel variances directly (feature vector $\mathbf{v}_{2m}$) are shown in Fig. 7(a1). A better segmentation is achieved if instead of the Euclidean metric one uses a weighted distance in which each component is standardized with respect to the current intraregion variance estimate (Fig. 7(b1)). The first component ($w_{1,2m}$) of our reduced feature set is shown in Fig. 7(c1), next to the corresponding segmentation obtained by simple thresholding (Fig. 7(d1)). By adding the second component $w_{2,2m}$ (Fig. 7(e1)) and applying the weighted K-means algorithm to this reduced feature set ($M = 2$), we obtain the segmentation map displayed in Fig. 7(f1). The corresponding results at a coarser scale ($2m = 8$) are shown in Fig. 7(a2–f2). Interestingly, the simple threshold segmentations ($M = 1$) are not much worse than the standard K-means results. A more accurate segmentation is achieved

if one uses a weighted distance criterion (Fig. 7(b1, b2)). However, the results obtained with the reduced feature sets (Fig. 7(f1, f2)) are essentially equivalent. The main advantage of our feature reduction method is a very substantial reduction in the amount of computation and a simplification of the segmentation process (especially for the case $K = 2$). Overall, the approach appears to be quite robust. The same type of behavior was observed over a variety of other testing conditions; these included different types of DWF's, different number features N, and a variety of texture classes.

The present approach to texture segmentation is one of the simplest conceivable for this particular class of texture features. In the current formulation, the multiscale structure of the feature extraction process is only exploited for feature reduction. Clustering is performed at a single scale of averaging, and the spatial resolution of the segmentation map therefore depends on the size of the smoothing window $(2m)$. This approach could be potentially improved by developing a true multiscale strategy to refine the location of the texture edges or by linking the decision process with some statistical model of the spatial region interactions (e.g., Gibbs distribution). There are already several extensions of the K-means algorithm that incorporate spatial interactions and that could be directly applied to our reduced feature set [33]–[35]; unfortunately, these tend to be extremely time-consuming. Another recent approach that appears to be very relevant uses a simple contour length constraint and a region-merging strategy [36].

V. Conclusion

One of the main points of this work has been to demonstrate that the wavelet transform provides an attractive tool for the characterization of texture properties. The wavelet transform has many properties (multiresolution representation, orthogonality, fast algorithms, etc.) that are relevant to this type of application. The main conclusions that can be drawn from the present analysis and from our classification experiments are as follows:

1) For feature extraction, an overcomplete wavelet analysis (DWF) is preferable to the standard subsampled wavelet decomposition. The DWF approach tends to decrease the variability of the estimated texture features (which improves classification performance). It also results in a texture characterization invariant under translation.
2) The multiresolution properties of the wavelet transform are beneficial for texture discrimination. In all cases, a multiscale feature extraction with two or three levels led to better results than a single resolution analysis. Moreover, the equivalence between the DWF for a single scale analysis and the method of local linear transforms (the performance of which is well documented in the literature) suggests that the present approach should outperform most traditional approaches to texture.
3) Increasing the number of vanishing moments (or regularity) of the underlying basis functions does not seem to have any real advantage for texture analysis and discrimination. In fact, the lower order transformations tend to perform surprisingly well, which is good news for practical applications. Another related observation is that the localization properties of the analysis filter bank seem to be more important than the properties of the basis functions themselves.
4) Finally, the DWF approach lends itself quite well to texture segmentation. This has been illustrated with a simple multicomponent algorithm that uses clustering in a reduced local feature space. The advantage of this procedure is to provide adequate segmentation at a much lower cost than an approach that operates on the full feature set.

Appendix
Cubic B-Spline Filtering

The discrete cubic B-spline kernel $b_m^3(k)$ is obtained by sampling a normalized version of the continuous cubic B-spline enlarged by a factor of m

$$b_m^3(k) = \frac{1}{m}\beta^3(k/m) \tag{A1}$$

where $\beta^3(x)$ is the symmetrical piecewise cubic function

$$\beta^3(x) := \begin{cases} 2/3 - x^2 + |x|^3/2, & 0 \le |x| < 1 \\ (2 - |x|)^3/6, & 1 \le |x| < 2 \\ 0, & 2 \le |x|. \end{cases} \tag{A2}$$

The z-transform of this kernel can be factorized as (cf. [30])

$$\begin{aligned} B_m^3(z) &= z^{2m} B_1^3(z) \left(\frac{1}{m} \sum_{k=0}^{m-1} z^{-k} \right)^4 \\ &= z^{2m} B_1^3(z) \left(\frac{1 - z^m}{m - m \cdot z^{-1}} \right)^4 \end{aligned} \tag{A3}$$

where z^{2m} is a shift factor, $B_1^3(z)$ denotes the B-spline sampled at the integers

$$B_1^3(z) = (z + 4 + z^{-1})/6, \tag{A4}$$

and where the term in parenthesis represents a moving average filter of length m. This factorization suggests a recursive implementation of b_m^3 based on the cascade of four moving average filters and the three-point FIR filter b_1^3.

The cubic spline kernel has some interesting scaling properties. In particular, its size can be doubled via the following convolution product:

$$B_{2m}^3(z) = H(z^m) B_m^3(z) \tag{A5}$$

where $H(z)$ is the fourth-order binomial filter

$$H(z) = \frac{z^2 + 4z + 6 + 4z^{-1} + z^{-2}}{16}. \tag{A6}$$

This result can be established easily by using (A3) and the fact that $\sum_{k=0}^{2m-1} z^{-k} = (1 + z^{-m}) \sum_{k=0}^{m-1} z^{-k}$. Unlike (3), the two-scale relation (A5) is valid when starting with any integer m (not just a power of two). It is a property that is specific to splines.

REFERENCES

[1] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, no. 6, pp. 610–621, Nov. 1973.

[2] P. C. Chen and T. Pavlidis, "Segmentation by texture using correlation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-5, pp. 64–69, Jan. 1983.

[3] R. L. Kashyap, R. Chellappa, and A. Khotanzad, "Texture classification using features derived from random field models," *Patt. Recogn. Lett.*, vol. 1, pp. 43–50, 1982.

[4] K. I. Laws, "Textured image segmentation," Ph.D. dissertation, Rept. 940, Image Processing Inst., Univ. of Southern Calif., 1980.

[5] M. Unser, "Local linear transforms for texture measurements," *Signal Processing*, vol. 11, no. 1, pp. 61–79, July 1986.

[6] R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, pp. 786–804, 1979.

[7] L. Van Gool, P. Dewaele, and A. Oosterlink, "Texture analysis Anno 1983," *Comput. Vision, Graph., Image Processing*, vol. 29, pp. 336–357, 1983.

[8] B. Julesz, "A theory of preattentive texture discrimination based on first-order statistics of textons," *Biol. Cybern.*, vol. 41, pp. 131–138, 1981.

[9] R. De Valois and K. De Valois, *Spatial Vision*. New York: Oxford, 1988.

[10] S. Marcelja, "Mathematical description of the responses of simple cortical cells," *J. Opt. Soc. Amer.*, vol. 70, no. 11, pp. 1297–1300, Nov. 1980.

[11] J. Beck, A. Sutter, and R. Ivry, "Spatial frequency channels and perceptual grouping in texture segregation," *Comput. Vision, Graph., Image Processing*, vol. 37, pp. 299–325, 1987.

[12] J. M. Coggins and A. K. Jain, "A spatial filtering approach to texture analysis," *Patt. Recogn. Lett.*, vol. 3, no. 3, pp. 195–203, 1985.

[13] M. Porat and Y. Y. Zeevi, "Localized texture processing in vision: Analysis and synthesis in Gaborian space," *IEEE Trans. Biomed. Eng.*, vol. 36, no. 1, pp. 115–129, 1989.

[14] M. R. Turner, "Texture discrimination by Gabor functions," *Biol. Cybern.*, vol. 55, pp. 71–82, 1986.

[15] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 12, no. 1, pp. 55–73, Jan. 1990.

[16] I. Daubechies, "Orthogonal bases of compactly supported wavelets," *Comm. Pure Appl. Math.*, vol. 41, pp. 909–996, 1988.

[17] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Mag.*, vol. 8, no. 4, pp. 11–38, Oct. 1991.

[18] S. G. Mallat, "Multiresolution approximations and wavelet orthogonal bases of $L^2(R)$," *Trans. Amer. Math. Soc.*, vol. 315, no. 1, pp. 69–87, 1989.

[19] ——, "A theory of multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, no. 7, pp. 674–693, 1989.

[20] T. Chang and C.-C. J. Kuo, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Trans. Image Processing*, vol. 2, no. 4, pp. 429–441, Oct. 1993.

[21] A. Laine and J. Fan, "Texture classification by wavelet packet signatures," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 15, no. 11, pp. 1186–1191, Nov. 1993.

[22] M. Unser, A. Aldroubi, and M. Eden, "A family of polynomial spline wavelet transforms," *Signal Processing*, vol. 30, no. 2, pp. 141–162, Jan. 1993.

[23] ——, "On the asymptotic convergence of B-spline wavelets to Gabor functions," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 864–872, Mar. 1992.

[24] Y. Meyer, *Ondelettes et Opèrateurs I : Ondelettes*. Paris: Hermann, 1990.

[25] O. Rioul, "A discrete-time multiresolution theory," *IEEE Trans. Signal Processing*, vol. 41, no. 8, pp. 2591–2606, Aug. 1993.

[26] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design," *IEEE Trans. Signal Processing*, vol. 40, no. 9, pp. 2207–2232, Sept. 1992.

[27] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia: Soc. Ind. Applied Math., 1992.

[28] R. J. Duffin and A. C. Schaeffer, "A class of nonharmonic Fourier series," *Trans. Amer. Math. Soc.*, vol. 72, pp. 314–366, 1952.

[29] M. Unser and M. Eden, "Multi-resolution feature extraction and selection for texture segmentation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, no. 7, pp. 717–728, July 1989.

[30] M. Unser, A. Aldroubi, and M. Eden, "Fast B-spline transforms for continuous image representation and interpolation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 13, no. 3, pp. 277–285, Mar. 1991.

[31] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Ann. Eugenics*, vol. 7, part II, pp. 179–188, 1936.

[32] G. B. Coleman and H. C. Andrews, "Image segmentation by clustering," *Proc. IEEE*, vol. 67, no. 5, pp. 773–785, May 1979.

[33] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-6, no. 6, pp. 712–741, Nov. 1984.

[34] J. Besag, "On the statistical analysis of dirty pictures," *J. Roy. Stat. Soc. B*, vol. 48, no. 3, pp. 256–302, 1986.

[35] S. Lakshmanan and H. Derin, "Simultaneous parameter estimation and segmentation of Gibbs random fields," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, no. 8, pp. 799–813, Aug. 1989.

[36] G. Koepfler, C. Lopez, and J. M. Morel, "A multiscale algorithm for image segmentation by variational method," *SIAM J. Numer. Anal.*, vol. 31, no. 1, pp. 282–299, Feb. 1994.

Michael Unser (M'88–SM'94) was born in Zug, Switzerland, April 9, 1958. He received the M.S. (summa cum laude) and Ph.D. degrees in electrical engineering in 1981 and 1984, respectively, from the Swiss Federal Institute of Technology in Lausanne, Switzerland.

He is currently a Visiting Scientist with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD, which he joined in 1985. He has also been affiliated with the French National Institutes of Health and Biomedical Research (INSERM) since April 1988. His research interests include the application of image processing and pattern recognition techniques to various biomedical problems, multiresolution algorithms, wavelet transforms, and the use of splines in signal processing.

Dr. Unser was awarded the Brown-Bowery Prize in 1984 for his work on texture analysis and automated inspection. He is the author of more than 50 published journal papers. He serves as Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING and the IEEE SIGNAL PROCESSING LETTERS and is a member of the Image and Multidimensional Signal Processing Committee of the IEEE Signal Processing Society. He is also on the editorial board of *Signal Processing* and *Pattern Recognition*.